

Short-Term Load Forecasting Using Time Pooling Deep Recurrent Neural Network

Elahe Khoshbakhti Vaygan

Faculty of ECE

Qom University of Technology

Qom, Iran

khoshbakhti.e@qut.ac.ir

Roozbeh Rajabi

Faculty of ECE

Qom University of Technology

Qom, Iran

rajabi@qut.ac.ir

Abouzar Estebarsari

School of the Built Environment and Architecture

London South Bank University

London, United Kingdom

estebasaa@lsbu.ac.uk

Abstract—Integration of renewable energy sources and emerging loads like electric vehicles to smart grids brings more uncertainty to the distribution system management. Demand Side Management (DSM) is one of the approaches to reduce the uncertainty. Some applications like Nonintrusive Load Monitoring (NILM) can support DSM, however they require accurate forecasting on high resolution data. This is challenging when it comes to single loads like one residential household due to its high volatility. In this paper, we review some of the existing Deep Learning-based methods and present our solution using Time Pooling Deep Recurrent Neural Network. The proposed method augments data using time pooling strategy and can overcome overfitting problems and model uncertainties of data more efficiently. Simulation and implementation results show that our method outperforms the existing algorithms in terms of RMSE and MAE metrics.

Index Terms—Load forecasting, Time pooling, Deep learning, Smart grids.

I. INTRODUCTION

One of the main objectives of distribution systems is to ensure continuity of service to final customers and make a balance between supply and demand. Therefore, it is crucial to estimate both generation and load in advance to manage system in operation more efficiently [1]. Load forecasting has been a key element in Distribution Management Systems (DMS) in operation and planning analyses as storing electric energy in large amount is very costly. An accurate estimation of load help match supply to avoid imbalances due to power deficit or surplus.

The measurements of electrical load over a long period of time not only show different values of loads over a periodical cycle (e.g. 24 hours), but also realize some trends of load variation in midterm or long-term horizons [2], [3]. Therefore, the load forecasting methods aim at predicting the load values in different time horizons using historic measurements, load trends analysis, experimental rules and mathematical models [4]. Some early load forecasting methods included Experimental Smoothing Models, Moving Average, and Auto Regression. The forecasting methods could be classified in three groups based on the techniques used: Gray Prediction Models, Statistical Analysis Models and Non-linear Intelligent Models [5], [6]. The most common practical load forecasting methods in traditional passive distribution networks are Auto

Regressive Integration Moving Average (ARIMA) and Support Vector Regression (SVR) [5], [7]. Integration of more electronic intelligent devices and communication technologies to distribution systems from one side and changing supply paradigms from bulk generation based models to distributed generation create o-called Smart Grids. In Smart Grids, there is a great opportunity for many advanced operation and analysis methods to be applied and utilized. These methods include Artificial Intelligence (AI) application like Machine Learning (ML) techniques and Deep Learning (DL) methods [8], [9], [10], [11].

In This paper, an AI method based on deep learning and pooling strategy techniques is proposed to increase the accuracy of load forecasting of the most volatile loads including single residential households. Similar to supervised multi-layer neural networks, the historic data is divided into two groups containing test and training sets, pooling test and pooling train, respectively. Before presenting our model, the complexity of load forecast in single residential customers and the parameters involved is briefed, and a literature review of ARIMA, SVR, Recurrent Neural Network (RNN) and Deep RNN is presented. The main contribution of this paper is to use time pooling strategy to enhance the performance of deep recurrent neural networks, overcoming the overfitting problem, and better modeling of the data.

The rest of the paper is organized as follows. Section II reviews the challenges regarding load forecasting and its applications. In this section previous methods used for short-term load forecasting also summarized. The proposed method using pooling strategy and deep learning is presented in III. Section IV describes the used dataset and experiments to evaluated the proposed TPRNN. Finally section V concludes the paper.

II. LITERATURE REVIEW

The residential appliances cover a wide range of devices with different characteristics, duty cycles, switching requirements, and power consumption. They are also different in terms of load types, from resistive types to inductive loads. This would make the aggregated profile very volatile with sharp increases or decreases of power in short periods of time. Moreover, the consumption behavior of residential customers

is not so predictable and there are always a high level of uncertainty. The uncertainty is due to different cultural and social consumption patterns, weather conditions which could be intermittent, and residual loads considered as noise. Since residential loads have a substantial share in total demand of distribution systems, the prediction of such stochastic loads is essential to manage the system. The importance of more accurate load forecasting is stressed when the system faces intermittent behavior of renewable energy sources (RES) widely integrated into the network.

Due to high volatility and uncertainty of single residential loads, we believe advanced ML-based methods should be applied to get a high accuracy in the results. DL techniques have been proposed in literature and in several projects as efficient methods, however the accuracy and performance of the applying DLs directly are arguable. In these methods, increase of deep layers is a fundamental approach to improve the accuracy of the results compare to Artificial Neural Networks (ANN), but in single residential load forecasting, the increase of DL layers exponentially increases the intrinsic parameters of the network in the training set. This would eventually grow the system noise which can affect the forecasting results. In summary, it is important to carefully model and adapt the DL network for particular types of loads and specific forecast time horizons in order to achieve desired accuracy and performance [12], [13]. Some parameters which may be used as extra input to the model include the economical and environmental conditions of the loads such as temperature, humidity, lighting, population growth, wind speed, economical situation, time and policies (Fig. 1).

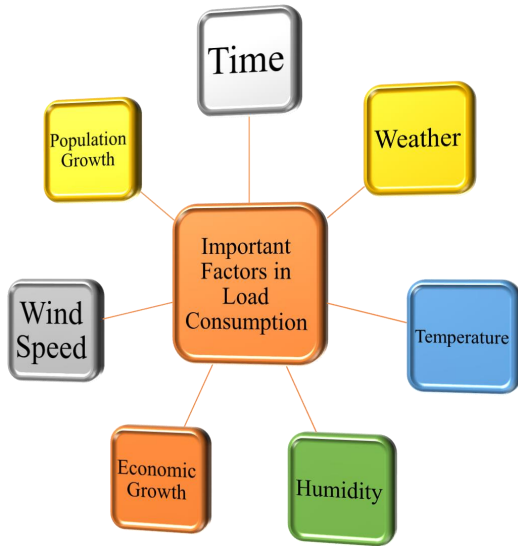


Fig. 1. Important factors having impact on load consumption.

In the following we briefly review 4 general methods to forecast loads in short-term: Auto Regressive Integration Moving Average (ARIMA), Support Vector Regression (SVR),

Recurrent Neural Network (RNN) and Deep Recurrent Neural Network (DRNN).

Auto Regressive Integration Moving Average

Time series models are the most common and popular methods to perform short-term forecasting. The Auto Regressive Integration Moving Average model is derived from Regressive Moving Average method which is time-series based method. In this model, time-series $y(t)$ is defined based on the previous values $y(t-1), y(t-2), \dots, y(t-p)$ and stochastic noise. The order of the process depends on the oldest previous value to which $y(t)$ is returned. This model is formulated as follow:

$$y(t) = \delta + \phi_1 y(t-1) + \dots + \phi_p y(t-p) + \varepsilon(t) \quad (1)$$

where $y(t-1), y(t-2), \dots, y(t-p)$ are the previous values of time series, $\varepsilon(t)$ is the stochastic noise, and δ is defined as follow [5]:

$$\delta \equiv \left(1 - \sum_{i=1}^p \phi_i\right) \mu \quad (2)$$

where μ is the average value of the process.

Support Vector Regression

SVR is based on Vapnik-chervonenkis theory and aim at expanding data sets to those not seen yet [7], [14]. SVM can be extended to be used as SVR. For this, a zone named ε -insensitivities defined around the function as ε -tube; then the SVR is formulated as follow [7]:

$$f(x) = \begin{pmatrix} x \\ 1 \end{pmatrix} \begin{pmatrix} W \\ b \end{pmatrix}^T = w^T + b \quad (3)$$

$x, w \in \mathbb{R}^{m+1}$

Recurrent Neural Network

The Recurrent Neural Network is based on Long Short-Term Memory (LSTM) methods which is able to exploit the interdependence in time-series over long periods of time. This makes the predictions of electrical loads more accurate. Considering $x = \{x_1, x_2, \dots, x_T\}$ as input time series, and $h = \{h_1, h_2, \dots, h_T\}$ and $y = \{y_1, y_2, \dots, y_T\}$ as output time series, which are created by updating the states of memory cell, LSTM decides what data to be removed from the cell gate [15]:

$$f(x) = \sigma(w_{fx}x_t + w_{fh}h_{t-1} + w_{fc}c_{t-1} + b_f) \quad (4)$$

In the next step, LSTM stores the new data as new cell state. this would be the input gate layer named i_t and formulated as the following:

$$i_t = \sigma(w_{ix}x_t + w_{ih}h_{t-1} + w_{ic}c_{t-1} + b_i) \quad (5)$$

A new vector is created, U_t , in order to store the new data that are going to be added as the new state to the cell. This vector is formed as the following equation:

$$U_t = g(w_{cx}x_t + w_{ch}h_{t-1} + b_c) \quad (6)$$

Then, the old cell state C_{t-1} is replaced by the new cell state c_t , using f_t and U_t :

$$c_t = U_t i_t + c_{t-1} f_t \quad (7)$$

The output of the output gate is represented by the following equation:

$$o_t = \sigma(w_{ox}x_t + w_{oh}h_{t-1} + w_{oc}c_{t-1} + b_o) \quad (8)$$

The new data, which is h_t , would be as follow:

$$h_t = o_t l(c_t) \quad (9)$$

And the final output would equal y_t :

$$y_t = k(w_{yh}h_t + b_y) \quad (10)$$

Where w_{ix} , w_{fx} , w_{ox} , and w_{cx} are input matrices; w_{ih} , w_{fh} , w_{oh} , and w_{ch} are recurrent weight matrices; and w_{yh} is the hidden output weight matrix. The matrices w_{ic} , w_{fc} , and w_{oc} are connection weight matrices, and the vectors b_i , b_f , b_o , b_c , and b_y are based on bias vector [15].

Deep Recurrent Neural Network

Deep Recurrent Neural Network (DRNN) has more layers in its network compare to RNN and has more benefits comparatively. Figure 2 shows the structure of DRNN and the functional process of a Deep RNN [12]. For a time-step, t , the parameters of the network neuron on the layer l updates their values according to the following equations [12]:

$$a_l^{(t)} = b_l + w_l h_l^{(t-1)} + U_l x^{(t)} \quad (11)$$

$$h_l^{(t)} = \text{activation function} \left(a_l^{(t)} \right) \text{ for } l = 1, 2, \dots, N \quad (12)$$

$$a_l^{(t)} = b_l + w_l h_l^{(t-1)} + U_l h_{l-1}^{(t)} \text{ for } l = 2, 3, \dots, N \quad (13)$$

$$y^{(t)} = b_N + w_N h_N^{(t-1)} + U_N h_N^{(t)} \quad (14)$$

$$L = \text{loss f} \left(y^{(t)}, y_{\text{target}}^{(t)} \right) \quad (15)$$

where, $x^{(t)}$ is the input of the time step t , $y^{(t)}$ is the predicted value, and $y_{\text{target}}^{(t)}$ is the output value. $h_l^{(t)}$ is the activation function at layer l of the network in time step t . $a_l^{(t)}$ is the real input value on layer l in time step t . The characteristics of the DRNN enables it to learn the uncertainties which are not very precedent.

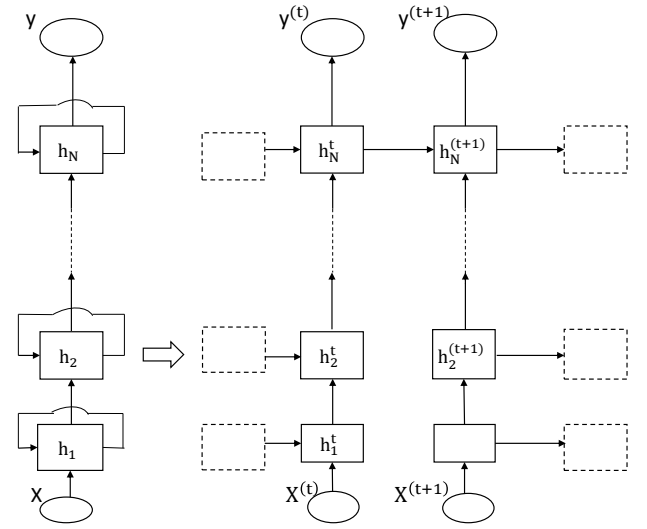


Fig. 2. Structure of DRNN.

III. PROPOSED METHOD

In this section the proposed methodology is presented. Pooling strategy [12] is used in the proposed method to tackle two main challenges in short-time load forecasting for single house residential applications. Overfitting is one of the major limits in deep learning-based load forecasting. Because of many layers in deep networks, training of these networks needs a large amount of data. Pooling stage can increase the amount of data greatly and avoid overfitting. The second challenge is uncertainty of load consumption data. Due to probabilistic nature of data, it is usually very difficult to model the data. Although some of these uncertainties are due to extrinsic factors such as weather conditions, diversity in training data can improve learning of the uncertainties. So, in the proposed method, pooling strategy is used to augment data and improve learning capabilities of the proposed method. A block diagram of the proposed method is illustrated in Fig. 3. The first step is loading data and data completion as discussed in IV-A.

After preprocessing stage, the data for each week is partitioned in a time-based manner. Total number of 10,080 measurements are available for each week with the resolution of one minute. This amount of data is divided to half day intervals ($N=24*60/2=720$) that will generate M number of groups ($M=14$). Each group will divided into training and test parts generating pools of data to train and test the model. In the next stage, deep recurrent neural network (DRNN) is trained to learn the load consumption profiles and finally the performance of the proposed method is evaluated using test dataset and metrics described in IV-C.

IV. EXPERIMENTS AND RESULTS

A. Dataset and Preprocessing

In this study, the individual household electric power consumption dataset is used to validate the proposed algorithm

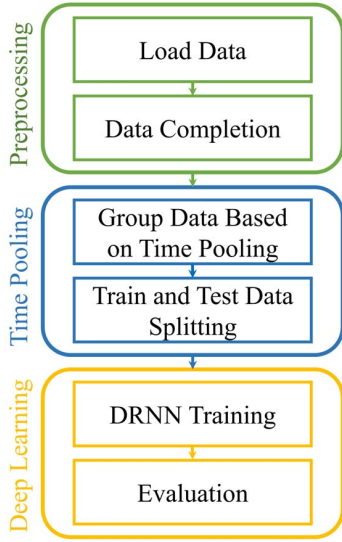


Fig. 3. The proposed time pooling DRNN method.

[16]¹. This dataset consists of 2,075,259 instance points that are measured over a period of almost 4 years. The global active power (GAP) in kW, date and time columns of this dataset are used in this paper. Preprocessing step based on averaging over available data is used to complete missing data. Fig. 4 shows GAP measurements of two consecutive weeks from the dataset exhibiting roughly similar pattern during a week. This similarity is exploited in the proposed method.

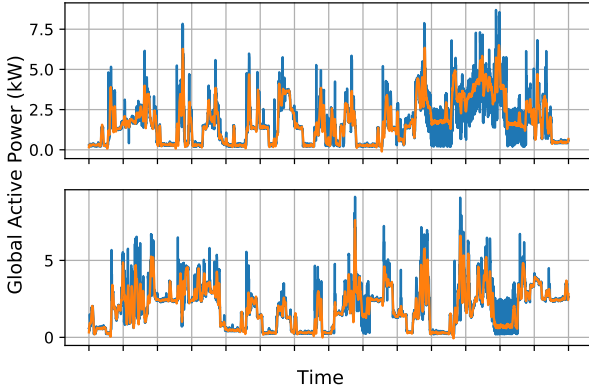


Fig. 4. Sample measurements from the dataset for two consecutive weeks starting from 2006-12-18 and ending to 2006-12-31.

B. Experiments

All experiments are done using Python programming language on a machine with 8 GB of RAM and Intel core i5-5200U up to 2.7 GHz CPU. The parameters are set as follows: N=720, M=14, training to test ratio: 67/23. Fig. 5 and

¹Available online: <https://archive.ics.uci.edu/ml/datasets/individual+household+electric+power+consumption>

6 show the training and test results of two methods DRNN and TPRNN (proposed method) respectively.

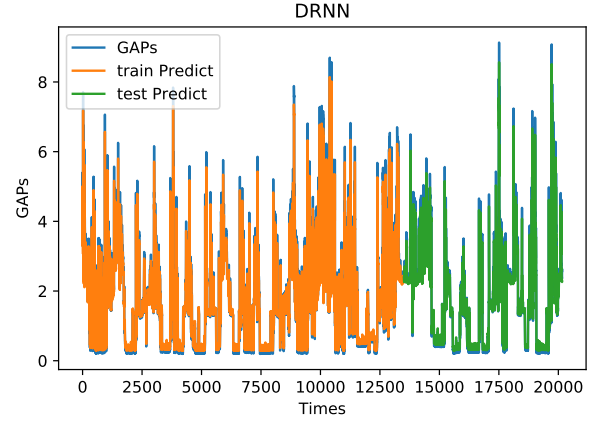


Fig. 5. Results of DRNN method for 2 weeks period.

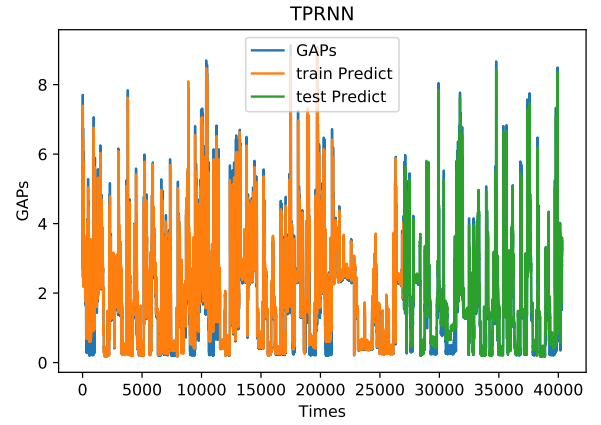


Fig. 6. Results of proposed method (TPRNN) for 4 weeks period.

C. Results and Discussion

To evaluate the proposed algorithm three widely used metrics are employed: Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). These metrics are defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (16)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (17)$$

where N is the total number of sample points. y_i and \hat{y}_i are the true and estimated values respectively.

Results are compared with different algorithms including SVR, ARIMA, RNN and DRNN in Table. I. The proposed method (TPRNN) exhibits better performance in comparison with other state-of-the-art methods.

TABLE I
COMPARISON RESULTS BASED ON RMSE AND MAE.

Method	RMSE	MAE
SVR	0.96	0.77
ARIMA	0.81	0.75
RNN	0.75	0.55
DRNN	0.39	0.20
TPRNN	0.37	0.19

V. CONCLUSION

Short-term load forecasting is an input or prerequisite for many applications in power systems, specially in smart grids. Depending on the application, different types of data are provided to run load forecasting algorithms. The accuracy of the forecasting is important for different applications too. In this paper, we presented a new short-term load forecasting method based on time pooling DRNN to predict the power demand of single residential customers for applications which require very short-term forecast. An application which could use this method is nonintrusive load monitoring (NILM). A high resolution historic data is used for training the model.

Load profiles demonstrate similarities in time. Based on these similarities, in this paper time-pooling strategy in conjunction with deep recurrent neural networks is used to solve the short-term load forecasting problem. Using pooling strategy has two fold advantages. Firstly it satisfies the need for large amount of data in deep learning methods and helps to model uncertainties in the dataset precisely. Results demonstrates that the proposed method can effectively forecast load consumption profile. Further research could be done including more precise time pooling using clustering algorithms to put similar power consumption patterns in a group.

REFERENCES

- [1] A. Garulli, S. Paoletti, and A. Vicino, "Models and techniques for electric load forecasting in the presence of demand response," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 3, pp. 1087–1097, 2014.
- [2] A. Estebarsari and R. Rajabi, "Single residential load forecasting using deep learning and image encoding techniques," *Electronics*, vol. 9, no. 1, p. 68, 2020.
- [3] R. Rajabi and A. Estebarsari, "Deep learning based forecasting of individual residential loads using recurrence plots," in *2019 IEEE Milan PowerTech*. IEEE, Conference Proceedings, pp. 1–5.
- [4] H. N. Rafsanjani, "Factors influencing the energy consumption of residential buildings: a review," in *Construction Research Congress 2016*, Conference Proceedings, pp. 1133–1142.
- [5] F. Mahia, A. R. Dey, M. A. Masud, and M. S. Mahmud, "Forecasting electricity consumption using ARIMA model," in *2019 International Conference on Sustainable Technologies for Industry 4.0 (STI)*. IEEE, Conference Proceedings, pp. 1–6.
- [6] A. Srivastava, A. S. Pandey, and D. Singh, "Notice of violation of ieee publication principles: Short-term load forecasting methods: A review," in *2016 International conference on emerging trends in electrical electronics & sustainable energy systems (ICETEESSES)*. IEEE, Conference Proceedings, pp. 130–138.
- [7] M. Awad and R. Khanna, *Support vector regression*. Springer, 2015, pp. 67–80.
- [8] T. Hossen, S. J. Plathottam, R. K. Angamuthu, P. Ranganathan, and H. Salehfar, "Short-term load forecasting using deep neural networks (DNN)," in *2017 North American Power Symposium (NAPS)*. IEEE, Conference Proceedings, pp. 1–6.
- [9] A. Almalaq and G. Edwards, "A review of deep learning methods applied on load forecasting," in *2017 16th IEEE international conference on machine learning and applications (ICMLA)*. IEEE, Conference Proceedings, pp. 511–516.
- [10] E. Busseti, I. Osband, and S. Wong, "Deep learning for time series modeling," *Technical report, Stanford University*, pp. 1–5, 2012.
- [11] H. Chen, C. A. Canizares, and A. Singh, "ANN-based short-term load forecasting in electricity markets," in *2001 IEEE power engineering society winter meeting. Conference proceedings (Cat. No. 01CH37194)*, vol. 2. IEEE, Conference Proceedings, pp. 411–415.
- [12] H. Shi, M. Xu, and R. Li, "Deep learning for household load forecasting—a novel pooling deep rnn," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5271–5280, 2017.
- [13] X. Sun, P. B. Luh, K. W. Cheung, W. Guan, L. D. Michel, S. S. Venkata, and M. T. Miller, "An efficient approach to short-term load forecasting at the distribution level," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2526–2537, 2016.
- [14] X. Sun, P. B. Luh, K. W. Cheung, W. Guan, L. D. Michel, S. Venkata, and M. T. Miller, "An efficient approach to short-term load forecasting at the distribution level," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2526–2537, 2015.
- [15] J. Zheng, C. Xu, Z. Zhang, and X. Li, "Electric load forecasting in smart grids using long-short-term-memory based recurrent neural network," in *2017 51st Annual Conference on Information Sciences and Systems (CISS)*. IEEE, Conference Proceedings, pp. 1–6.
- [16] G. Hebrail and A. Berard, "Individual household electric power consumption data set," *É. d. France, Ed., ed: UCI Machine Learning Repository*, 2012.